

STIC-Biotech/ChemLib

82698

From: Ibrahim, Medina A.
Sent: Wednesday, December 18, 2002 6:49 PM
To: STIC-Biotech/ChemLib
Subject: 09/538, 396

RECEIVED

DEC 19 2002

STIC-BIOTECH/CHEM. DIVISION
(STIC)

Please search the following:

1. oligo search of at least 30 contiguous bases of SEQ ID NO:1.
2. residue of at least 20 contiguous amino acids of SEQ ID NO:2. Thanks

Medina A. Ibrahim
Patent Examiner
GAU:1638
CM1-9A12
(703)306-5822

Point of Contact:
Toby Post
Technical info. Specialist
CM1 6A04
703-308-3534

Searcher: _____
Phone: _____
Location: _____
Date Picked Up: 12/20
Date Completed: 12/16
Searcher Prep/Review: _____
Clerical: _____
Online time: _____

TYPE OF SEARCH:
NA Sequences: _____
AA Sequences: _____
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST (where applic.)
STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: _____
WWW/Internet: _____
Other (specify): _____

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 26, 2002, 05:24:41 : Search time 125 Seconds
(without alignments)
14598.440 Million cell updates/sec

Title: US-09-538-396-1
Perfect score: 4492
Sequence: 1 aattcgcaagatgagtc.....atctagatcaaaaaaaaaa 4492

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 363474 seqs, 203117208 residues
Word size : 30

Total number of hits satisfying chosen parameters: 0

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : Published_Applications_NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description

No matches found

Search completed: December 26, 2002, 08:43:45
Job time : 127 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compen Ltd.

OM protein - protein search, using sw model

Run on: December 26, 2002, 08:43:51 ; Search time 27 Seconds
(without alignments)
1434.095 Million cell updates/sec

Title: US-09-538-396-2
Perfect score: 1316
Sequence: 1 MSTVDKMLIKIGRSFDPDK.....YKSKDENQHSIISEQIFD 1316

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 262574 seqs, 29422922 residues

Word size : 20

Total number of hits satisfying chosen parameters: 2

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Database : Issued Patents_AA:*
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/6C.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/6D.COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	31	2.4	1312	2	US-08-592-126-148
2	31	2.4	1312	2	US-08-687-080-51

ALIGNMENTS

RESULT 1
US-08-592-126-148
Sequence 148, Application US/08592126
Patent No. 5821091
GENERAL INFORMATION:
APPLICANT: Gregory Dolganov
TITLE OF INVENTION: Transcripts Encoding Immunomodulatory
TITLE OF INVENTION: Polypeptides
NUMBER OF SEQUENCES: 151
CORRESPONDENCE ADDRESS:
ADDRESSER: Dehlinger & Associates
STREET: 350 Cambridge Avenue, Suite 250
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/592,126
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Sholtz, Charles K.
REGISTRATION NUMBER: 38,615
REFERENCE/DOCKET NUMBER: 4600-0111
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 148:
SEQUENCE CHARACTERISTICS:
LENGTH: 1312 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: Rad50.pro-translation of SEQ ID NO:54
US-08-592-126-148
Query Match 2.4%; Score 31; DB 2; Length 1312;
Best Local Similarity 100.0%; Pred. No. 4.8e-20;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1207 MGRGCSAGQVYLAIIIRLALAEFTCLNCIGI 1237
Db 1197 MGRGCSAGQVYLAIIIRLALAEFTCLNCIGI 1227

RESULT 2
US-08-687-080-51
Sequence 51, Application US/08687080
Patent No. 5965427
GENERAL INFORMATION:
APPLICANT: Gregory Dolganov
TITLE OF INVENTION: Human RAD50 Gene and Methods of Use Thereof
NUMBER OF SEQUENCES: 175
CORRESPONDENCE ADDRESS:
ADDRESSER: Dehlinger & Associates
STREET: 350 Cambridge Avenue, Suite 250
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/687,080
FILING DATE: 17-JUL-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/592,126
FILING DATE: 26-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Sholtz, Charles K.
REGISTRATION NUMBER: 38,615
REFERENCE/DOCKET NUMBER: 4600-0111.30
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 51:
SEQUENCE CHARACTERISTICS:
LENGTH: 1312 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 26, 2002, 08:45:21 : Search time 39 Seconds
(without alignments)
584.241 Million cell updates/sec

Title: US-09-538-396-2
Perfect score: 1316
Sequence: 1 MSTVDKMKLIKIRSPDPNK.....YRVSKDENQHSIESQELFD 1316

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 109717 seqs, 17314136 residues

Word size : 20

Total number of hits satisfying chosen parameters: 0

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :

Published_Applications_AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
No matches found					

Search completed: December 26, 2002, 08:52:12
Job time : 39 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 26, 2002, 08:41:21 ; Search time 46 Seconds
(without alignments)
2750.280 Million cell updates/sec

Title: US-09-538-396-2

Perfect score: 1316

Sequence: 1 MSTVDMKLKIGIRSFDPDNK.....YRVSKDENQHSIIIESQEIFD 1316

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 283224 seqs, 96134422 residues

Word size : 20

Total number of hits satisfying chosen parameters: 2

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :

1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	49	3.7	1292	DB4727	probable RAD50 DNA
2	42	3.2	1312	T30845	probable DNA repair

ALIGNMENTS

RESULT 1

DB4727 Probable RAD50 DNA repair protein [Imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 01-Mar-2002

C:Accession: DB4727

R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.

euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.

Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A:Reference number: A84420; MUID:20083487; PMID:10617197

A:Accession: DB4727

A:Status: Preliminary

A:Molecule type: DNA

A:Residues: 1-1292 <STO>

A:Cross-references: GB:AE002093; NID:g4263721; PIDN:ADD15407.1; GSPDB:GN00139

C:Genetics:

A:Gene: At2g31970

A:Map position: 2

C:Superfamily: RAD50 protein

Query Match 3.7%; Score 49; DB 2; Length 1292;
Best Local Similarity 100.0%; Pred. No. 6.6e-40;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 152 LENVIFVHODESNWPLQDPSTLKKKFDIFSATRTKALEVTKLHKDQ 200
DB 152 LENVIFVHODESNWPLQDPSTLKKKFDIFSATRTKALEVTKLHKDQ 200

RESULT 2

T30845

probable DNA repair protein RAD50 - mouse

C:Species: Mus musculus (house mouse)

C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 17-Mar-2000

C:Accession: T30845

R:Kim, K.K.; Daud, A.I.; Wong, S.C.; Pajak, L.; Tsai, S.C.; Wang, H.; Henzel, W.J.; F

J. Biol. Chem. 271, 29255-29264, 1996

A:Title: Mouse RAD50 has limited epitopic homology to p53 and is expressed in the adu

A:Reference number: 220899; MUID:97067183; PMID:8910585

A:Accession: T30845

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1312 <KIM>

A:Cross-references: EMBL:U66887; NID:g1575574; PID:g1575575; PIDN:AAC52894.1

C:Genetics:

A:Gene: RAD50

A:Map position: 11

C:Superfamily: RAD50 protein

C:Keywords: DNA repair

Query Match 3.2%; Score 42; DB 2; Length 1312;
Best Local Similarity 100.0%; Pred. No. 7.2e-33;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1207 MRGCSAGQKVLASLIRLALAEFTCLNGCILAIDEPPTNLD 1248

DB 1197 MRGCSAGQKVLASLIRLALAEFTCLNGCILAIDEPPTNLD 1238

Search completed: December 26, 2002, 08:50:24
Job time : 46 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 26, 2002, 07:47:17 : Search time 25 Seconds
(without alignments)
2183.314 Million cell updates/sec

Title: US-09-538-396-2
Perfect score: 1316
Sequence: 1 MSTVDKMLIKIRSFDPDNK.....YRVSKDENQHSHIESQEIFD 1316

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 112892 seqs, 41476328 residues

Word size : 20

Total number of hits satisfying chosen parameters: 0

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	ID	Description
No matches found			

Search completed: December 26, 2002, 08:47:34
Job time : 25 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 26, 2002, 05:15:21 : Search time 104 Seconds
(without alignments)
13246.077 Million cell updates/sec

Title: US-09-538-396-1
Perfect score: 4492
Sequence: 1 aatcgcgcagcagtgatcc.....atctagcatcaaaaaaaaa 4492

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 441362 seqs, 15338381 residues

Word size : 30

Total number of hits satisfying chosen parameters: 0

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/Backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result	Query	Score	Match	Length	ID	Description

No matches found						

Search completed: December 26, 2002, 08:41:16
Job time : 105 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 26, 2002, 08:43:51 ; Search time 27 Seconds
(without alignments)
1434.095 Million cell updates/sec

Title: US-09-538-396-2

Perfect score: 1316
Sequence: 1 MSYVDKMLKINGISFDPDK.....YRYSKDNQHSIESQELFD 1316

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 262574 seqs, 29422922 residues

Word size: 20

Total number of hits satisfying chosen parameters: 2

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database:

Issued Patents_AA:*
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCJUS.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/Backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	31	2.4	1312	2	US-08-592-126-148
2	31	2.4	1312	2	US-08-687-080-51

ALIGNMENTS

RESULT 1
US-08-592-126-148

Sequence 148, Application US/08592126

Patent No. 5821491

GENERAL INFORMATION:

APPLICANT: Gregory Dolganov

TITLE OF INVENTION: Tinscripts Encoding Immunomodulatory

NUMBER OF SEQUENCES: Polypeptides

NUMBER OF SEQUENCES: 151

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Dehlinger & Associates

STREET: 350 Cambridge Avenue, Suite 250

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94306

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/592,126

FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Sholtz, Charles K.

REGISTRATION NUMBER: 38,615

REFERENCE/DOCKET NUMBER: 4600-0111

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 324-0860

TELEFAX: (415) 324-0960

INFORMATION FOR SEQ ID NO: 148:

SEQUENCE CHARACTERISTICS:

LENGTH: 1312 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: NO

ORIGINAL SOURCE:

INDIVIDUAL ISOLATE: Rad50, pro-translation of SEQ ID NO:54

US-08-592-126-148

Query Match

Best Local Similarity 2.4%; Score 31; DB 2; Length 1312;

Best Local Similarity 100.0%; Pred. No. 4.8e-20;

Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1207 MRGRCSAGQVLAIIIRLAETFCINCGI 1237

DB 1197 MRGRCSAGQVLAIIIRLAETFCINCGI 1227

RESULT 2
US-08-687-080-51

Sequence 51, Application US/08687080

Patent No. 5965427

GENERAL INFORMATION:

APPLICANT: Gregory Dolganov

TITLE OF INVENTION: Human RAD50 Gene and Methods of Use Thereof

NUMBER OF SEQUENCES: 175

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Dehlinger & Associates

STREET: 350 Cambridge Avenue, Suite 250

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94306

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/687,080

FILING DATE: 17-JUL-1996

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/592,126

FILING DATE: 26-JAN-1996

ATTORNEY/AGENT INFORMATION:

NAME: Sholtz, Charles K.

REGISTRATION NUMBER: 38,615

REFERENCE/DOCKET NUMBER: 4600-0111.30

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 324-0860

TELEFAX: (415) 324-0960

INFORMATION FOR SEQ ID NO: 51:

SEQUENCE CHARACTERISTICS:

LENGTH: 1312 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

;
 ; HYPOTHETICAL: NO
 ; ANTI-SENSE: NO
 ; ORIGINAL SOURCE:
 ; INDIVIDUAL ISOLATE: TRANS. OF RAD50 CDNA (SEQ. 54), NT.
 ; INDIVIDUAL ISOLATE: 389 TO 4524
 ; US-08-687-080-51

Query Match 2.4%; Score 31; DB 2; Length 1312;
 Best Local Similarity 100.0%; Pred. No. 4.8e-20;
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1207 /MRGRCSAGQKVLASLIRLALAEFFCLNCGI 1237
 Db 1197 /MRGRCSAGQKVLASLIRLALAEFFCLNCGI 1227

Search completed: December 26, 2002, 08:51:12
 Job time : 27 secs

XX	AAC66915;
AC	
XX	27-MAR-2001 (first entry)
DT	
XX	
DE	Maize RAD50 coding sequence.
XX	
KW	Maize; RAD50; meiotic recombination; DNA repair; chromosome 4p; recombination efficiency; transformation efficiency; ss.
XX	
OS	Zea mays.
PN	WO20068404-A1.
PD	16-NOV-2000.
PF	25-APR-2000; 2000WO-US1086.
PR	05-MAY-1999; 99US-0132575.
PA	(PION-) PIONEER HI-BRED INT INC.
PI	MahaJan PB, Shi J;
PT	Nucleic acids encoding Rad50 polypeptides, useful for transgenic plant production and for producing antibodies that are useful for modifying gene expression -
PS	Claim 1; Page 65-70; 76pp; English.
XX	
CC	The present sequence is the coding sequence of the maize RAD50 gene, found on maize chromosome 4p. RAD50 is involved in meiotic recombination and DNA repair, and the gene and its protein can be used to produce transgenic plants whose expression of the gene may be regulated. This can be useful in the regulation of transformation and recombination efficiency in plants.
CC	
SO	Sequence 4492 BP; 1537 A; 831 C; 1061 G; 1063 T; 0 other;
Query Match	100.0%; Score 4492; DB 22; Length 4492;
Best Local Similarity	100.0%; Pred. No. 0;
Matches 4492; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
OY	1 AATTCGGCAGCATGATCGATTAGCACCCATAGCCGTACAAAACCCTAAGAACCCTAAC 60
DB	1 AATTCGGCAGCATGATCGATTAGCACCCATAGCCGTACAAAACCCTAAGAACCCTAAC 60
OY	61 CGGTACAAAACCTTA AAAACCTTAAGCGCGCTGAAGACTCCAAAAAGCGGATTTTCTC 120
DB	61 CGGTACAAAACCTTA AAAACCTTAAGCGCGCTGAAGACTCCAAAAAGCGGATTTTCTC 120
OY	121 CTCGACGCCGCCCTCTTCTCTTCCAACTGTTTTGCATCACTAGAGCGTATGAAT 180
DB	121 CTCGACGCCGCCCTCTTCTCTTCCAACTGTTTTGCATCACTAGAGCGTATGAAT 180
OY	181 AGAAGTGATAGGAGATATGATCCGCAATCTAGGTTTGGGCAATGCTCTGGCCAGAC 240
DB	181 AGAAGTGATAGGAGATATGATCCGCAATCTAGGTTTGGGCAATGCTCTGGCCAGAC 240
OY	241 TGATTCGAGAGTCAAGTCTGAGAGGAGGACAATTGGGGCTCTGGGGCCAAGTAGAGACC 300
DB	241 TGATTCGAGAGTCAAGTCTGAGAGGAGGACAATTGGGGCTCTGGGGCCAAGTAGAGACC 300
OY	301 GTTACCAAGATGCTGATCAAGGGGATTCGGAGCTTGATCCGGACAATAAGACGTCATC 360
DB	301 GTTACCAAGATGCTGATCAAGGGGATTCGGAGCTTGATCCGGACAATAAGACGTCATC 360
OY	361 ACCCTTTCAAAGCGCTACACCTCATGCTGTTGGCCCCCAACGGTGTCTGCAAGACGATC 420
DB	361 ACCCTTTCAAAGCGCTACACCTCATGCTGTTGGCCCCCAACGGTGTCTGCAAGACGATC 420


```

|||||
Db 2581 GAGAAAACCTGATCAACATTTGGCGATGAAAGTCAGAAAGCGCAGCGCATTTGATCAT 2640
|||
Oy 2641 CTTTGGGCTGCTTCTGGCCATGTCATTAATGACAGGGATGAGTGGAGGCTTTATTACA 2700
|||
Db 2641 CTTTGGGCTGCTTCTGGCCATGTCATTAATGACAGGGATGAGTGGAGGCTTTATTACA 2700
|||
Oy 2701 CCCACTGATCTATTGACAGCATGTACATGAAATTCACAGCTAGTCACAAAGTAGAA 2760
|||
Db 2701 CCCACTGATCTATTGACAGCATGTACATGAAATTCACAGCTAGTCACAAAGTAGAA 2760
|||
Oy 2761 GATCTGAAATATGACACTGATTTAGTGGCCGAGGTGTCAAAGCTTTGGAGAAATTCAA 2820
|||
Db 2761 GATCTGAAATATGACACTGATTTAGTGGCCGAGGTGTCAAAGCTTTGGAGAAATTCAA 2820
|||
Oy 2821 CTGAGAGCTGAATCTTCTGACAGAAACAGAGACATGATTTGTCAGAGTGGATGATCTT 2880
|||
Db 2821 CTGAGAGCTGAATCTTCTGACAGAAACAGAGACATGATTTGTCAGAGTGGATGATCTT 2880
|||
Oy 2881 AGAGATCAACATAGATCTAAATGAAATGATATGTCAGTGTCTCAGGTGAGATGCGACAT 2940
|||
Db 2881 AGAGATCAACATAGATCTAAATGAAATGATATGTCAGTGTCTCAGGTGAGATGCGACAT 2940
|||
Oy 2941 GCTCGGAGAGAGAAAGTGAAGCTTCTAGCATATTGGAGAAAGTTCACAAAGCTGAAGG 3000
|||
Db 2941 GCTCGGAGAGAGAAAGTGAAGCTTCTAGCATATTGGAGAAAGTTCACAAAGCTGAAGG 3000
|||
Oy 3001 GAATGGGCTTCTAGCTGAGAGAAAGACAACTGATTTGAGAAAGAGCTTTTAGAA 3060
|||
Db 3001 GAATGGGCTTCTAGCTGAGAGAAAGACAACTGATTTGAGAAAGAGCTTTTAGAA 3060
|||
Oy 3061 GAGTCTCTGATCCATTTCTCCAAAGAGAAAGAGAGCTTGTCCAAAGATATATGCTTTG 3120
|||
Db 3061 GAGTCTCTGATCCATTTCTCCAAAGAGAAAGAGAGCTTGTCCAAAGATATATGCTTTG 3120
|||
Oy 3121 AAGCAAAAGCTGATGAAAGATATCATCAGCTTGCAGAAAGAGAAAGAGAGTTCAGCAA 3180
|||
Db 3121 AAGCAAAAGCTGATGAAAGATATCATCAGCTTGCAGAAAGAGAGAAAGAGAGTTCAGCAA 3180
|||
Oy 3181 GAACCTGATGCTCTGGAAGACTTATATGAGATGAAAGGCTTACTGGATTCAGAGAA 3240
|||
Db 3181 GAACCTGATGCTCTGGAAGACTTATATGAGATGAAAGGCTTACTGGATTCAGAGAA 3240
|||
Oy 3241 AAGCAAAAGCTTAAAGAAATTCAGAGAGAGAGCTTCTTGGCATCTCAGTATACAGAGT 3300
|||
Db 3241 AAGCAAAAGCTTAAAGAAATTCAGAGAGAGAGCTTCTTGGCATCTCAGTATACAGAGT 3300
|||
Oy 3301 TGCATGGCAAAACAGCAAGAAATATCAGCTGAGTTAAACAGAGCAAGAAAGCTGACG 3360
|||
Db 3301 TGCATGGCAAAACAGCAAGAAATATCAGCTGAGTTAAACAGAGCAAGAAAGCTGACG 3360
|||
Oy 3361 GGCCAGGGGCGCATGGAAGAAACATGATGACAAATCTCAGTACAGAGAAACAAAGGCT 3420
|||
Db 3361 GGCCAGGGGCGCATGGAAGAAACATGATGACAAATCTCAGTACAGAGAAACAAAGGCT 3420
|||
Oy 3421 GATGTGGAACAATCTACTGCTGATATAGATCACTTGAAGAAAGGCTGCTTCAATAGT 3480
|||
Db 3421 GATGTGGAACAATCTACTGCTGATATAGATCACTTGAAGAAAGGCTGCTTCAATAGT 3480
|||
Oy 3481 AGCTTGTCTGCTATAGAGCTGATCTGAACCGCATTTCTCAAGAAAGAGAGCTTAAT 3540
|||
Db 3481 AGCTTGTCTGCTATAGAGCTGATCTGAACCGCATTTCTCAAGAAAGAGAGCTTAAT 3540
|||
Oy 3541 TCAGAAATTTAAAGGTGGGAAGGAACACTTCTGTTATCAAGTATATTTCAAGAC 3600
|||
Db 3541 TCAGAAATTTAAAGGTGGGAAGGAACACTTCTGTTATCAAGTATATTTCAAGAC 3600
|||
Oy 3601 AAACAAGAGCTTAACTGTACAGTACAAAGATATGAGAAAGGATATCTAAATCAATT 3660
|||
Db 3601 AAACAAGAGCTTAACTGTACAGTACAAAGATATGAGAAAGGATATCTAAATCAATT 3660
|||
Oy 3661 CTCACACTTAAAGCAACTGAAATGCAAAACAGAGCTTGGACAGATTTTACTGCTTTA 3720
|||

```

```

Db 3661 CTCACACTTAAAGCAACTGAAATGGCAAAACAAGAGCTTGGACAGATATTTACTGCTTTA 3720
|||
Oy 3721 GACAAGGCTCTTATGCGGCTTCCACAGCATGAAGATGAGAGATTAATTAATCAAG 3780
|||
Db 3721 GACAAGGCTCTTATGCGGCTTCCACAGCATGAAGATGAGAGATTAATTAATCAAG 3780
|||
Oy 3781 GAACGTGGCAACAGACATTAAGAGGCGGAGATTTGATTCATATAGCATTAATTCGAT 3840
|||
Db 3781 GAACGTGGCAACAGACATTAAGAGGCGGAGATTTGATTCATATAGCATTAATTCGAT 3840
|||
Oy 3841 TCTGAGGCTGCGCACTGATCATACAGTACAGTACCGGCTGTTTATGCAAACTGATGAGCT 3900
|||
Db 3841 TCTGAGGCTGCGCACTGATCATACAGTACAGTACCGGCTGTTTATGCAAACTGATGAGCT 3900
|||
Oy 3901 GAGCTGGAAGTCGAGGCGGCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3960
|||
Db 3901 GAGCTGGAAGTCGAGGCGGCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3960
|||
Oy 3961 AGACTAGCACTTGGCGAAACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4020
|||
Db 3961 AGACTAGCACTTGGCGAAACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4020
|||
Oy 4021 ACTAGCAATCTAGATGGGCAAAATGCAAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4080
|||
Db 4021 ACTAGCAATCTAGATGGGCAAAATGCAAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4080
|||
Oy 4081 GAAGCAGAGAAAGGCGAGAGAACTTCCAGTGTGATTTATATCAATCAATGAGAGATTT 4140
|||
Db 4081 GAAGCAGAGAAAGGCGAGAGAACTTCCAGTGTGATTTATATCAATCAATGAGAGATTT 4140
|||
Oy 4141 GGCCATCTTATGGGTCAAAAGGCAAGCTGCTGAGAGTACTATGAGTCTCCAAAGATGAG 4200
|||
Db 4141 GGCCATCTTATGGGTCAAAAGGCAAGCTGCTGAGAGTACTATGAGTCTCCAAAGATGAG 4200
|||
Oy 4201 AACCAAGCAGCATATTTGAATCCAAAGATTTTGAATGAGTCTTGAAGGCTG 4260
|||
Db 4201 AACCAAGCAGCATATTTGAATCCAAAGATTTTGAATGAGTCTTGAAGGCTG 4260
|||
Oy 4261 TAGCAGCACTGCTTGTGCTAGTGCAGATCCAGTATTTATGCAAGTACTGGTGCAGAG 4320
|||
Db 4261 TAGCAGCACTGCTTGTGCTAGTGCAGATCCAGTATTTATGCAAGTACTGGTGCAGAG 4320
|||
Oy 4321 CAATGTTAAACAAGCTTAAAGAGCTGTTAGAGGCTGTTAGAGGCTGTTAGAGGCTGTTAG 4380
|||
Db 4321 CAATGTTAAACAAGCTTAAAGAGCTGTTAGAGGCTGTTAGAGGCTGTTAGAGGCTGTTAG 4380
|||
Oy 4381 CCTTGGTGTGATATCTTATTAATCTGGCCAGAGATGAAATGTGTGCACTGGGTGAT 4440
|||
Db 4381 CCTTGGTGTGATATCTTATTAATCTGGCCAGAGATGAAATGTGTGCACTGGGTGAT 4440
|||
Oy 4441 GGATGTTTCAAGCATCAATGAATGTTTCAAAATCTAGCATCAAAAGAGAG 4492
|||
Db 4441 GGATGTTTCAAGCATCAATGAATGTTTCAAAATCTAGCATCAAAAGAGAG 4492
|||

```

Search completed: December 26, 2002, 05:28:18
 Job time : 362 secs

ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
Sorghum bicolor					
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Sorghum.					
1 (bases 1 to 512)					
Reid, S.P., Cordonnier-Pratt, M.-M., Gingle, A. and Pratt, L.H.					
An EST database from Sorghum: developing embryos					
Unpublished (2000)					
Contact: Cordonnier-Pratt MM					
Laboratory for Genomics and Bioinformatics					
The University of Georgia, Department of Plant Biology					
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA					
Tel: 706 542 1860					
Fax: 706 583 0210					
Email: mpratt@uga.edu					
Sequences have been trimmed to exclude polyA, vector and regions below Phred quality 16. The threshold for highest quality sequence is 20.					
Seq primer: JEN REV					
High quality sequence stop: 383					
POLYA-No.					
Location/Qualifiers					
1..512					
/organism="Sorghum bicolor"					
/db_xref="taxon:4558"					
/clone_lib="Embryo 1 (EM1)"					
/note="Organ: Embryos germinated for 24 hr; Vector: Bluescript II from Lambda Zap II; Site 1: XhoI; Site 2: EcoRI; The library was made from poly-A RNA in the cloning vector lambda Zap II. Clones to be sequenced were prepared by mass excision."					
BASE COUNT					
ORIGIN					
102 a					
168 c					
157 g					
85 t					
Query Match					
Best Local Similarity					
Matches 149; Conservative					
2.2%; Score 99; DB 12; Length 512;					
99.3%; Pred. No. 6.5e-37;					
0; Mismatches 1; Indels 0; Gaps 0;					
Oy 3726					
GGCTCTTATGGCGTTCACAGCATGAGATGAGAGATAAATAATCAAGAACT					
358					
GGCTCTTATGGCGTTCACAGCATGAGATGAGAGATAAATAATCAAGAACT					
417					
Oy 3786					
GTGGCAACAGCATTCACAGGCCAGGATTTGATTACATAAGCAATAATTCGATTCTGA					
3845					
GTGGCAACAGCATTCACAGGCCAGGATTTGATTACATAAGCAATAATTCGATTCTGA					
477					
Oy 3846					
GGGTCTGTCGACATCGATCATACAGCTACCG					
3875					
GGGTCTGTCGACATCGATCATACAGCTACCG					
478					
GGGTCTGTCGACATCGATCATACAGCTACCG					
507					
RESULT 3					
BG649231					
LOCUS					
DEFINITION					
EM1_78_H06.g1_A002 Embryo 1 (EM1) Sorghum bicolor cDNA, mRNA					
sequence.					
ACCESSION					
BG649231					
VERSION					
BG649231.1					
KEYWORDS					
EST.					
SOURCE					
ORGANISM					
Sorghum bicolor					
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Sorghum.					
1 (bases 1 to 484)					
Reid, S.P., Cordonnier-Pratt, M.-M., Gingle, A. and Pratt, L.H.					
An EST database from Sorghum: developing embryos					
Unpublished (2000)					
Contact: Cordonnier-Pratt MM					
Laboratory for Genomics and Bioinformatics					
The University of Georgia, Department of Plant Biology					
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA					
Tel: 706 542 1860					

FEATURES	source
<p> Fax: 706 583 0210 Email: impratt@uga.edu Sequences have been trimmed to exclude PolyA, vector and regions below Phred quality 16. The threshold for highest quality sequence is 20. Seq primer: PolyTMax High quality sequence start: 7 High quality sequence stop: 484 POLYA-No. </p>	<p> Location/Qualifiers 1. .484 /organism="Sorghum bicolor" /db_xref="taxon:4558" /clone_1lb="Embryo 1 (EM1)" /note="Organ: Embryos germinated for 24 hr; Vector: pBluescript II from lambda zap II; Site_1: XhoI; Site_2: PstI; The library was made from polyA RNA in the cloning vector lambda ZAP II. Clones to be sequenced were prepared by mass excision." </p>
<p> BASE COUNT ORIGIN Query Match Best Local Similarity 100.0%; Pred. No. 3.1e-20; Matches 65; Conservative 0; Mismatches 0; Indels 0; Gaps 0; </p>	<p> 145 a 92 c 113 g 134 t </p>
<p> QY 4066 CTGTTGAGATATATGGAAGCCAGAAAGGCGAGAACTTCAGTTGATGTAATCACT 4125 Db 47 CTGTGAGATATATGGAAGCCAGAAAGGCGAGAACTTCAGTTGATGTAATCACT 106 </p>	<p> 1.4%; Score 65; DB 12; Length 484; 100.0%; Pred. No. 3.1e-20; </p>
<p> QY 4126 CATGA 4130 Db 107 CATGA 111 </p>	<p> 416 bp mRNA linear EST 04-JAN-2002 P1C1.50.G12.G1.A002 Pathogen-Infected compatible 1 (P1C1) Sorghum bicolor cDNA, mRNA sequence. BM330353 BM330353.1 GI:18069490 EST. </p>
<p> RESULT 4 BM330353 LOCUS DEFINITION ACCESSION VERSION KEYWORDS SOURCE ORGANISM </p>	<p> Sorghum. Sorghum bicolor Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Sorghum. 1 (bases 1 to 416) Cordonnier-Pratt,M.-M., Gingle,A., Fang,G.C., Dean,R., Wing,R., Sudman,M. and Pratt,L.H. An EST database from Sorghum: plants infected with a compatible pathogen </p>
<p> REFERENCE AUTHORS TITLE JOURNAL COMMENT </p>	<p> Unpublished (2002) Contact: Cordonnier-Pratt MM Laboratory for Genomics and Bioinformatics The University of Georgia, Department of Plant Biology Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA Tel: 706 542 1860 Fax: 706 583 0210 Email: impratt@uga.edu Sequences have been trimmed to exclude PolyA, vector, and regions below Phred quality 16. The threshold for highest quality sequence is 20. Three-prime sequences, which are obtained with PolyTmix or T7 sequencing primer, are presented as the reverse complement. Seq primer: T7 High quality sequence stop: 308 POLYA-yes. </p>
<p> FEATURES source </p>	<p> Location/Qualifiers 1. .416 /organism="Sorghum bicolor" /cultivar="BRx623" /db_xref="taxon:4558" </p>

/clone_1lb="Pathogen-infected compatible 1 (PIC1)"
 /tissue_type="Leaves"
 /dev_stage="4-week-old seedlings infected with
 Colletotrichum graminicola"
 /note="Vector: pBluescript II SK(-) from Lambda Zap II;
 Site_1: XhoI; Site_2: EcoRI; Four-week-old sorghum
 seedlings were sprayed with spore suspension prepared from
 3-week-old FRM421; a sorghum isolate of the anthracnose
 pathogen Colletotrichum graminicola. Inoculated plants
 were kept in a 25 C dark growth chamber with 100% relative
 humidity for 24 hr, followed by 12/12 hr of light/dark
 cycle at 25 C with 90% relative humidity for another 24
 hr. All leaves were harvested and quick frozen with liquid
 nitrogen and stored in a -80 C freezer. The library was
 made from poly-A RNA in the cloning vector Lambda Zap II.
 Clones to be sequenced were prepared by mass excision.
 WARNING: While most or all ESTs are expected to derive
 from the host plant, no effort was made to eliminate ESTs
 deriving from the pathogen."

BASE COUNT 121 a 75 c 100 g 120 t
 ORIGIN

Query Match 1.0%; Score 47; DB 13; Length 416;
 Best Local Similarity 100.0%; Pred. No. 2.1e-11;
 Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4153 GGTCAAGGACGCTGCTGAGAGTCTTCGAGTCCAGAGATGA 4199
 ||||||||||||||||||||||||||||||||||||||||||||
 Db 86 GGTCAAGGACGCTGCTGAGAGTCTTCGAGTCCAGAGATGA 132

RESULT 5
 BI416347 549 bp mRNA linear EST 14-AUG-2001
 LOCUS LjNESTd3r Lotus japonicus nodule library, mature and immature
 DEFINITION nodules Lotus japonicus cDNA 5', mRNA sequence.
 ACCESSION BI416347
 VERSION BI416347.1 GI:15186602
 KEYWORDS EST.
 SOURCE Lotus japonicus.
 ORGANISM Lotus japonicus
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Lotese;
 Lotus.
 1 (bases 1 to 549)
 Freund, S., Stougaard, J. and Urdavari, M.
 Lotus japonicus root nodule ESTs: a tool for functional genomics
 Unpublished (2000)
 CONTACT: Urdavari MK
 Molecular Plant Nutrition
 Max Planck Institute of Molecular Plant Physiology
 Am Muehlenberg 1, 14476 Golm, Germany
 Fax: 49 331 567 8250
 Email: urdavari@mplp-golm.mpg.de
 Seg primer: 77
 High quality sequence stop: 549.
 Location/Qualifiers
 1. 549
 /organism="Lotus japonicus"
 /cultivar="Gifu (B-129)"
 /db_xref="taxon:34305"
 /clone_1lb="Lotus japonicus nodule library, mature and
 immature nodules"
 /note="Organ: Nodule; Vector: pBluescript SK-; Site_1:
 EcoRI; Site_2: XhoI; The library was prepared using mRNA
 extracted from a mixture of mature (pink) and immature
 (white) nodules of Lotus japonicus ecotype 'Gifu'. Nodules
 were induced by, and contained Mesorhizobium loti strain
 N2P2235."

BASE COUNT 178 a 110 c 123 g 138 t
 ORIGIN

Query Match 0.8%; Score 35; DB 13; Length 549;
 Best Local Similarity 100.0%; Pred. No. 1.6e-05;
 Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3662 TCACGTTAGACAACTGAAATGCAACAGAGAC 3696
 ||||||||||||||||||||||||||||||||||||||||||||
 Db 157 TCACGTTAGACAACTGAAATGCAACAGAGAC 191

RESULT 6
 BE443510 458 bp mRNA linear EST 25-JUL-2000
 LOCUS WHE1115.D09.H17Z5 wheat etiolated seedling root normalized cDNA
 DEFINITION library Triticum aestivum cDNA clone WHE1115.D09.H17, mRNA
 sequence.

ACCESSION BE443510
 VERSION BE443510.1 GI:9443044
 KEYWORDS EST.
 SOURCE bread wheat.
 ORGANISM Triticum aestivum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
 Triticeae; Triticum.
 1 (bases 1 to 458)
 Anderson, O.D., Chao, S., Choi, D.W., Close, T.J., Fenton, R.D., Han
 'P.S.', Hsia, C.C., Kang, Y., Lazo, G.R., Miller, R., Nguyen, H.T.,
 Rausch, C.J., Seaton, C.L., Tong, J.C., and Zhang, D.
 The structure and function of the expressed portion of the wheat
 genomes - Normalized root cDNA library
 Unpublished (2000)
 CONTACT: Olin Anderson
 US Department of Agriculture, Agriculture Research Service, Pacific
 West Area, Western Regional Research Center
 800 Buchanan Street, Albany, CA 94710, USA
 Tel: 5105585773
 Fax: 5105585818
 Email: oanderson@pv.usda.gov
 Sequence have been trimmed to remove vector sequence and low
 quality sequence with phred score less than 20
 Seg primer: Stratagene SK primer.
 Location/Qualifiers
 1. 458
 /organism="Triticum aestivum"
 /cultivar="Chinese Spring"
 /db_xref="taxon:4565"
 /clone_1lb="WHE1115.D09.H17"
 /clone_1lb="wheat etiolated seedling root normalized cDNA
 library"
 /tissue_type="Root"
 /dev_stage="Five day old etiolated seedling"
 /lab_host="E. coli DH10B"
 /note="Vector: Lambda Uni-ZAP XR, excised phagemid
 pBluescript SK; Site_1: EcoRI; Site_2: XhoI; Seeds were
 surface-sterilized, germinated and grown aseptically in
 the dark at room temperature on filter paper with water,
 nystatin and cefotaxime in covered crystallization
 dishes. Roots were harvested. The tissue, total RNA, and
 poly(A) RNA were prepared. A cDNA library was made in the
 T7 close lab (Choi, Close, Fenton) at the University of
 California, Riverside. The cDNA clones were in vivo
 excised to give pBluescript phagemids before
 normalization was carried out. The mass excision of
 phagemid library and normalization were done in HT Nguyen
 lab by D. Zhang at Texas Tech University. Normalization
 protocol used was that of Soares. Plasmid DNA
 preparations and DNA sequencing were performed in the OD
 Anderson lab (all other authors)."

BASE COUNT 173 a 88 c 94 g 103 t
 ORIGIN

Query Match 0.7%; Score 30; DB 10; Length 458;
 Best Local Similarity 100.0%; Pred. No. 0.0047;
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Thu Dec 26 09:44:36 2002

us-09-538-396-1.0130.rst

Page 4

OY 1345 GGAAGCTCCAGCGAGAAGCTGATGCTCAC 1374
|||||
Db 291 GGAAGCTCCAGCGAGAAGCTGATGCTCAC 320

Search completed: December 26, 2002, 08:39:20
Job time : 3880 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model
Run on: December 26, 2002, 08:39:26 ; Search time 82 Seconds
(without alignments)
3306.805 Million cell updates/sec

Title: US-09-538-396-2
Perfect score: 1316
Sequence: 1 MSTRVDMKILKIGRSFDPDNK.....YRVSKDENOHSHIESQEIFD 1316

Scoring table:
Gapop 60.0 , Gapext 60.0

Searched: 671580 seqs, 206047115 residues

Word size: 20

Total number of hits satisfying chosen parameters: 6

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database:

SPTREMBL_21:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rhodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriophage:*
- 17: sp_archaeal:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	49	3.7	1316	10	Q9SL02
2	42	3.2	1312	11	Q9JIL8
3	42	3.2	1312	11	P70388
4	31	2.4	1173	4	Q9UP86
5	31	2.4	1312	4	Q92878
6	31	2.4	1318	4	O43254

ALIGNMENTS

RESULT 1
Q9SL02 PRELIMINARY; PRT; 1316 AA.
AC Q9SL02: Q9M6P9;
DT 01-MAY-2000 (TREMBlrel. 13, Created)

01-JUN-2002 (TREMBlrel. 21, Last sequence update)
01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Putative RAD50 DNA repair protein (DNA repair-recombination protein).
DE AR2G31970 OR RAD50.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV, COLUMBIA;
RA Lin X., Kaul S., Shea T.P., Fujita C.Y., Shen M., Vanaken S.E.,
RA Barnstead M.E., Mason T.M., Bowman C.L., Rinning C.M., Benito M.-I.,
RA Carrera A.J., Greasy T.H., Buell C.R., Town C.D., Nielsen W.C.,
RA Fraser C.M., Venter J.C.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.

[2]
RN SEQUENCE FROM N.A.
RC STRAIN=CV, COLUMBIA;
RX MEDLINE=21097002; PubMed=11169180;
RA Galleo M.E., Jeanneau M., Granier F., Bouchez D., Bechtold N.,
RA White C.I.;
RT Disruption of the Arabidopsis RAD50 gene leads to plant sterility and
RT MMS sensitivity.
RT Plant J. 25:31-41(2001).
DR EMBL: AC006223; AAD15407.2;
DR EMBL: AF168748; AAF36810.1;
DR InterPro: IPR003439; ABC_transporter.
DR InterPro: IPR004584; Rad50.
DR TIGRFRMS; TIGR00506; rad50; 1.
DR TIGRFRMS; TIGR00506; rad50; 1.
SQ SEQUENCE 1316 AA; 152814 MW; 89DC4FBCA39B0E8 CRC64;

Query Match 3.7%; Score 49; DB 10; Length 1316;
Best Local Similarity 100.0%; Pred. No. 6.2e-39;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 152 LENVIFVHODESNMPLDPSSTLKRRFDIFSARRYTKALEVYIKKKHQ 200
DB 152 LENVIFVHODESNMPLDPSSTLKRRFDIFSARRYTKALEVYIKKKHQ 200

RESULT 2
Q9JIL8 PRELIMINARY; PRT; 1312 AA.
AC Q9JIL8;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Rad50.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Granulata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20368553; PubMed=10908350;
RA Lanson N.A., Jr., Egeland D.B., Royals B.A., Claycomb W.C.;
RT The MRE11-NBS1-RAD50 pathway is perturbed in SV40 large T antigen-
RT immortalized At-1, At-2 and Hs-1 cardiomyocytes.
RL Nucleic Acids Res. 28:2882-2892(2000).
DR EMBL: AF218576; AAF91229.1;
DR InterPro: IPR003439; ABC_transporter.
DR InterPro: IPR001395; Aldo_ket-red.
DR InterPro: IPR004584; Rad50.
DR TIGRFRMS; TIGR00606; rad50; 1.
DR PROSITE; PS00063; ALDO-KETO_REDUCTASE_3; UNKNOWN.1.
SQ SEQUENCE 1312 AA; 153783 MW; F13C041BD2C05932 CRC64;

Query Match 3.2%; Score 42; DB 11; Length 1312;
Best Local Similarity 100.0%; Pred. No. 5.3e-32;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1207 MRGCSAGQKVLASLIIRLALAEFTCLNCIILADEPTNLD 1248
 ||||||||||||||||||||||||||||||||||||||||
 Db 1197 MRGCSAGQKVLASLIIRLALAEFTCLNCIILADEPTNLD 1238

RESULT 3

ID P70388 PRELIMINARY; PRT; 1312 AA.
 AC P70388;
 DT 01-FEB-1997 (TREMBLrel. 02, Created)
 DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE RAD50.
 GN RAD50 OR RAD50.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-97067183; PubMed-8910585;
 RA Kim K.K., Daud A.I., Wong S.C., Pajak L., Tsai S.C., Wang H.,
 Hanzel W.J., Field L.J.;
 RT "Mouse RAD50 has limited epitopic homology to p53 and is expressed in
 the adult myocardium";
 RL J. Biol. Chem. 271:29255-29264(1996).
 DR EMBL: D66887; AAC52894.1; -
 DR MGI: 109292; Rad50.
 DR InterPro: IPR003439; ABC_transport.
 DR InterPro: IPR001395; Aldo_ket_red.
 DR TIGRFAMs: TIGR00606; rad50.1.
 DR PROSITE: PS00063; ALDO_KETO_REDUCTASE_3; UNKNOWN.1.
 SQ SEQUENCE 1312 AA; 153487 MW; 4AF9AF9AD9ELD7A2 CRC64;

Query Match

Best Local Similarity 3.2%; Score 42; DB 11; Length 1312;
 Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1207 MRGCSAGQKVLASLIIRLALAEFTCLNCIILADEPTNLD 1248
 ||||||||||||||||||||||||||||||||||||||||
 Db 1197 MRGCSAGQKVLASLIIRLALAEFTCLNCIILADEPTNLD 1238

RESULT 4

ID Q90P86 PRELIMINARY; PRT; 1173 AA.
 AC Q90P86;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Truncated Rad50 protein.
 GN RAD50.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-99346140; PubMed-10415333;
 RA Kim K.K., Shin B.A., Seo K.H., Kim P.N., Koh J.T., Kim J.H.,
 Park B.R.;
 RT "Molecular cloning and characterization of splice variants of human
 RAD50 gene";
 RL Gene 235:59-67(1999).
 DR EMBL: AF057300; AAD50326.1; -
 DR InterPro: IPR003439; ABC_transport.
 DR InterPro: IPR001395; Aldo_ket_red.
 DR TIGRFAMs: TIGR00606; rad50.1.
 DR PROSITE: PS00063; ALDO_KETO_REDUCTASE_3; UNKNOWN.1.
 SQ SEQUENCE 1173 AA; 138432 MW; D6734A4EAD898AAE CRC64;

Query Match 2.4%; Score 31; DB 4; Length 1173;
 Best Local Similarity 100.0%; Pred. No. 3.8e-21;
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1207 MRGCSAGQKVLASLIIRLALAEFTCLNCI 1237
 ||||||||||||||||||||||||||||||||||||
 Db 1058 MRGCSAGQKVLASLIIRLALAEFTCLNCI 1088

RESULT 5

ID Q92878 PRELIMINARY; PRT; 1312 AA.
 AC Q92878;
 DT 01-FEB-1997 (TREMBLrel. 02, Created)
 DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE RAD50.
 GN RAD50.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-99346140; PubMed-10415333;
 RA Bressan D.A., Petrini J.H.J.;
 RA Dolganov G.M., Maser R.S., Novikov A., Tosto L., Chong S.,
 Mol. Cell. Biol. 16:0-0(0).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-99346140; PubMed-10415333;
 RA Kim K.K., Shin B.A., Seo K.H., Kim P.N., Koh J.T., Kim J.H.,
 Park B.R.;
 RT "Molecular cloning and characterization of splice variants of human
 RAD50 gene";
 RL Gene 235:59-67(1999).
 DR EMBL: U63139; AAB07119.1; -
 DR EMBL: AF057299; AAD50325.1; -
 DR InterPro: IPR003439; ABC_transport.
 DR InterPro: IPR001395; Aldo_ket_red.
 DR TIGRFAMs: TIGR00606; rad50.1.
 DR PROSITE: PS00063; ALDO_KETO_REDUCTASE_3; UNKNOWN.1.
 SQ SEQUENCE 1312 AA; 153891 MW; 1F208C3817FC411C CRC64;

Query Match

Best Local Similarity 2.4%; Score 31; DB 4; Length 1312;
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1207 MRGCSAGQKVLASLIIRLALAEFTCLNCI 1237
 ||||||||||||||||||||||||||||||||||||
 Db 1197 MRGCSAGQKVLASLIIRLALAEFTCLNCI 1227

RESULT 6

ID Q43254 PRELIMINARY; PRT; 1318 AA.
 AC Q43254;
 DT 01-JUN-1998 (TREMBLrel. 06, Created)
 DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE RAD50 homologue hsrRAD50.
 GN RAD50.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-TESTIS;
 RA Offenberg H.H.;
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL: Z75311; CAA99729.1; -

DR InterPro; IPR003439; ABC_transport.
 DR InterPro; IPR001395; Aldo/Ket_red.
 DR InterPro; IPR004584; Rad50.
 DR TIGRFAMs; TIGR00606; rad50; 1.
 DR PROSITE; PS00063; ALDOXETO_REDUCTASE_3; UNKNOWN_1.
 SO SEQUENCE 1318 AA; 154410 MW; B4EB9EEEB8105822 CRC64;

Query Match 2.4%; Score 31; DB 4; Length 1318;
 Best Local Similarity 100.0%; Pred. No. 4.2e-21;
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1207 MRGRCSAGQKVLASLIIRLALAEFCINCGI 1237
 |||||
 Db 1203 MRGRCSAGQKVLASLIIRLALAEFCINCGI 1233

Search completed: December 26, 2002, 08:49:18
 Job time : 83 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 26, 2002, 05:28:26 ; Search time 68 seconds
(without alignments)
2578.791 Million cell updates/sec

Title: US-09-538-396-2

Perfect score: 1316
Sequence: 1 MSTVDMKLIKIRSFDPDNK.....YRVSKDENQHSIIIESQEIFD 1316

Scoring table: -0.100 Gapop:60.0, Gapext:60.0

Searched: 908470 seqs, 133250620 residues

Word size: 20

Total number of hits satisfying chosen parameters: 3

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Database: A_Geneseq_101002.*

```
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1316	100.0	1316	22	Maize RAD50. Zea
2	31	2.4	1312	18	Human RAD50. Homo
3	31	2.4	1312	19	Human homologue of

ALIGNMENTS

RESULT 1
AAB27248

ID	Protein	Length	Score	Query Match	Length	ID	Protein	Length	Score	Query Match	Length
AC	AAB27248	1316	100.0	100.0	1316	AC	AAB27248	1316	100.0	100.0	1316
XX	27-MAR-2001 (first entry)	1316	100.0	100.0	1316	XX	27-MAR-2001 (first entry)	1316	100.0	100.0	1316
DE	Maize RAD50.	1316	100.0	100.0	1316	DE	Maize RAD50.	1316	100.0	100.0	1316
XX	Maize; RAD50; meiotic recombination; DNA repair; chromosome 4p; recombination efficiency; transformation efficiency.	1316	100.0	100.0	1316	XX	Maize; RAD50; meiotic recombination; DNA repair; chromosome 4p; recombination efficiency; transformation efficiency.	1316	100.0	100.0	1316
OS	Zea mays.	1316	100.0	100.0	1316	OS	Zea mays.	1316	100.0	100.0	1316
XX	WO200068404-A1.	1316	100.0	100.0	1316	XX	WO200068404-A1.	1316	100.0	100.0	1316
PD	16-NOV-2000.	1316	100.0	100.0	1316	PD	16-NOV-2000.	1316	100.0	100.0	1316
PF	25-APR-2000; 2000MO-US11086.	1316	100.0	100.0	1316	PF	25-APR-2000; 2000MO-US11086.	1316	100.0	100.0	1316
XX	05-MAY-1999; 99US-0132575.	1316	100.0	100.0	1316	XX	05-MAY-1999; 99US-0132575.	1316	100.0	100.0	1316
PA	(PION-) PIONEER HI-BRED INT INC.	1316	100.0	100.0	1316	PA	(PION-) PIONEER HI-BRED INT INC.	1316	100.0	100.0	1316
PI	Mahajan PB, Shi J;	1316	100.0	100.0	1316	PI	Mahajan PB, Shi J;	1316	100.0	100.0	1316
XX	WPI, 2001-007402/01.	1316	100.0	100.0	1316	XX	WPI, 2001-007402/01.	1316	100.0	100.0	1316
DR	N-PSDB; AAC66915.	1316	100.0	100.0	1316	DR	N-PSDB; AAC66915.	1316	100.0	100.0	1316
XX	Nucleic acids encoding Rad50 polypeptides, useful for transgenic plant production and for producing antibodies that are useful for modifying gene expression.	1316	100.0	100.0	1316	XX	Nucleic acids encoding Rad50 polypeptides, useful for transgenic plant production and for producing antibodies that are useful for modifying gene expression.	1316	100.0	100.0	1316
PT	Claim 11; Page 61-62; 76pp; English.	1316	100.0	100.0	1316	PT	Claim 11; Page 61-62; 76pp; English.	1316	100.0	100.0	1316
CC	The present sequence is the protein sequence of maize RAD50, the gene for which is found on maize chromosome 4p. RAD50 is involved in meiotic recombination and DNA repair, and the protein and its gene can be used to produce transgenic plants whose expression of the gene may be regulated. This can be useful in the regulation of transformation and recombination efficiency in plants.	1316	100.0	100.0	1316	CC	The present sequence is the protein sequence of maize RAD50, the gene for which is found on maize chromosome 4p. RAD50 is involved in meiotic recombination and DNA repair, and the protein and its gene can be used to produce transgenic plants whose expression of the gene may be regulated. This can be useful in the regulation of transformation and recombination efficiency in plants.	1316	100.0	100.0	1316
XX	Sequence 1316 AA;	1316	100.0	100.0	1316	XX	Sequence 1316 AA;	1316	100.0	100.0	1316
QY	Query Match	1316	100.0	100.0	1316	QY	Query Match	1316	100.0	100.0	1316
DB	Best Local Similarity 100.0%; Pred. No. 0;	1316	100.0	100.0	1316	DB	Best Local Similarity 100.0%; Pred. No. 0;	1316	100.0	100.0	1316
DB	Matches 1316; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	1316	100.0	100.0	1316	DB	Matches 1316; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	1316	100.0	100.0	1316
QY	1 MSTVDMKLIKIRSFDPDNKNTTFFKPLTLVGPAGAKTTIIECLKSTGELPNSR 60	1316	100.0	100.0	1316	QY	1 MSTVDMKLIKIRSFDPDNKNTTFFKPLTLVGPAGAKTTIIECLKSTGELPNSR 60	1316	100.0	100.0	1316
DB	1 MSTVDMKLIKIRSFDPDNKNTTFFKPLTLVGPAGAKTTIIECLKSTGELPNSR 60	1316	100.0	100.0	1316	DB	1 MSTVDMKLIKIRSFDPDNKNTTFFKPLTLVGPAGAKTTIIECLKSTGELPNSR 60	1316	100.0	100.0	1316
QY	61 SGHTFVHDPKVAEETKQIKLRFKTAGKDVVCIRSPQLOKASKMEFAIESVLTQI 120	1316	100.0	100.0	1316	QY	61 SGHTFVHDPKVAEETKQIKLRFKTAGKDVVCIRSPQLOKASKMEFAIESVLTQI 120	1316	100.0	100.0	1316
DB	61 SGHTFVHDPKVAEETKQIKLRFKTAGKDVVCIRSPQLOKASKMEFAIESVLTQI 120	1316	100.0	100.0	1316	DB	61 SGHTFVHDPKVAEETKQIKLRFKTAGKDVVCIRSPQLOKASKMEFAIESVLTQI 120	1316	100.0	100.0	1316
QY	121 NPHTGKVCISRYCADMDREIPALMGVSAVLENYFVHODESNMPLDPSFLAKKFDI 180	1316	100.0	100.0	1316	QY	121 NPHTGKVCISRYCADMDREIPALMGVSAVLENYFVHODESNMPLDPSFLAKKFDI 180	1316	100.0	100.0	1316
DB	121 NPHTGKVCISRYCADMDREIPALMGVSAVLENYFVHODESNMPLDPSFLAKKFDI 180	1316	100.0	100.0	1316	DB	121 NPHTGKVCISRYCADMDREIPALMGVSAVLENYFVHODESNMPLDPSFLAKKFDI 180	1316	100.0	100.0	1316
QY	181 FSATRYTKALEVYIKLHAKQMOEIKFRKLENTQVVKQAKRLKRENTAODDEKSDASKS 240	1316	100.0	100.0	1316	QY	181 FSATRYTKALEVYIKLHAKQMOEIKFRKLENTQVVKQAKRLKRENTAODDEKSDASKS 240	1316	100.0	100.0	1316
DB	181 FSATRYTKALEVYIKLHAKQMOEIKFRKLENTQVVKQAKRLKRENTAODDEKSDASKS 240	1316	100.0	100.0	1316	DB	181 FSATRYTKALEVYIKLHAKQMOEIKFRKLENTQVVKQAKRLKRENTAODDEKSDASKS 240	1316	100.0	100.0	1316
QY	241 QMEQLKEKICGERELTOMETSIDELRRLOGOIDIKATERSTLLTQOHEKLAALSEENED 300	1316	100.0	100.0	1316	QY	241 QMEQLKEKICGERELTOMETSIDELRRLOGOIDIKATERSTLLTQOHEKLAALSEENED 300	1316	100.0	100.0	1316
DB	241 QMEQLKEKICGERELTOMETSIDELRRLOGOIDIKATERSTLLTQOHEKLAALSEENED 300	1316	100.0	100.0	1316	DB	241 QMEQLKEKICGERELTOMETSIDELRRLOGOIDIKATERSTLLTQOHEKLAALSEENED 300	1316	100.0	100.0	1316
QY	301 TDEELMEMOTKFEERIALLETETKISKLYRMDDEASYSVLSKONSELTHEIGKLAEDA 360	1316	100.0	100.0	1316	QY	301 TDEELMEMOTKFEERIALLETETKISKLYRMDDEASYSVLSKONSELTHEIGKLAEDA 360	1316	100.0	100.0	1316
DB	301 TDEELMEMOTKFEERIALLETETKISKLYRMDDEASYSVLSKONSELTHEIGKLAEDA 360	1316	100.0	100.0	1316	DB	301 TDEELMEMOTKFEERIALLETETKISKLYRMDDEASYSVLSKONSELTHEIGKLAEDA 360	1316	100.0	100.0	1316
QY	361 HTMKHERDSIDIKINICTKNNLGPVPHPTNDVYANLITRIARLSSLENDLDDKKSYE 420	1316	100.0	100.0	1316	QY	361 HTMKHERDSIDIKINICTKNNLGPVPHPTNDVYANLITRIARLSSLENDLDDKKSYE 420	1316	100.0	100.0	1316
DB	361 HTMKHERDSIDIKINICTKNNLGPVPHPTNDVYANLITRIARLSSLENDLDDKKSYE 420	1316	100.0	100.0	1316	DB	361 HTMKHERDSIDIKINICTKNNLGPVPHPTNDVYANLITRIARLSSLENDLDDKKSYE 420	1316	100.0	100.0	1316

```

Db 361 HTMKEROSDKINICTKNNLGPVPHPTNDVAMLTNRKARLSLENDLLDKKKSNE 420
OY 421 DQDVLAKHKLINAYSEVDGQIOKISMSGILRRDRDKEKERDAAEVLSKFLSRI 480
OY 422 DQDVLAKHKLINAYSEVDGQIOKISMSGILRRDRDKEKERDAAEVLSKFLSRI 480
OY 481 DEBERMQLIEVERKTLALABRDYDSIIISOKRETVYSLEOKIKVLAERKILNNADERYK 540
OY 481 DEBERMQLIEVERKTLALABRDYDSIIISOKRETVYSLEOKIKVLAERKILNNADERYK 540
OY 481 DEBERMQLIEVERKTLALABRDYDSIIISOKRETVYSLEOKIKVLAERKILNNADERYK 540
OY 541 LGKKALSSSKDKLEIYNEHKOKIKVLRGNRPPEKMKKEIINAFNPVDEKEYNELS 600
OY 541 LGKKALSSSKDKLEIYNEHKOKIKVLRGNRPPEKMKKEIINAFNPVDEKEYNELS 600
OY 601 KSOEABOLKFTOSKYTDAREQLTKLRMDADARRFLDSKLSIIQISANVDFPVYLD 660
OY 601 KSOEABOLKFTOSKYTDAREQLTKLRMDADARRFLDSKLSIIQISANVDFPVYLD 660
OY 661 AMNRDQKRLNFPANGMREMLAPFEHLARKNHVCCCEAFPPDEDEPVKKQRQNS 720
OY 661 AMNRDQKRLNFPANGMREMLAPFEHLARKNHVCCCEAFPPDEDEPVKKQRQNS 720
OY 721 TAERSKALAESSNAEALFQQLDKRTIYDAYVYKVEETIPLAEKLNQHLADESOKAQA 780
OY 721 TAERSKALAESSNAEALFQQLDKRTIYDAYVYKVEETIPLAEKLNQHLADESOKAQA 780
OY 781 FDDLGLVAVOQDRDAVEALLQPTDITDRHVEHIOQLVEVEDLEVALDSGRGVKSL 840
OY 781 FDDLGLVAVOQDRDAVEALLQPTDITDRHVEHIOQLVEVEDLEVALDSGRGVKSL 840
OY 841 EIOLELNFORTBDTLIVEVDLROHRLNEDSSAQVWNAAREKVAASSILEPFOK 900
OY 841 EIOLELNFORTBDTLIVEVDLROHRLNEDSSAQVWNAAREKVAASSILEPFOK 900
OY 901 SEEPVLLAEKREOLYERKILLESIDPLSKERESLLOEYNALOKIDEYHOLAERKE 960
OY 901 SEEPVLLAEKREOLYERKILLESIDPLSKERESLLOEYNALOKIDEYHOLAERKE 960
OY 961 FQOELDALGRNLNMIKGYLDSKNEKLEKQGRHVLSHOLSQOSMAQOHSIAELNKSKE 1020
OY 961 FQOELDALGRNLNMIKGYLDSKNEKLEKQGRHVLSHOLSQOSMAQOHSIAELNKSKE 1020
OY 1021 LLOGOGQLKRNIDNLTAKRYKADVEQLFRDIESLEERLISIGLSAIEADLKRHSOKE 1080
OY 1021 LLOGOGQLKRNIDNLTAKRYKADVEQLFRDIESLEERLISIGLSAIEADLKRHSOKE 1080
OY 1081 RANSEFNWOGTISVYOSNISKHOKELKSOYKDIERYNOFOLKTTEKANDLDTRY 1140
OY 1081 RANSEFNWOGTISVYOSNISKHOKELKSOYKDIERYNOFOLKTTEKANDLDTRY 1140
OY 1141 TALDKALMBFMSKMEENKIKELMOOTYRGODIDYISINSDEGAGTSYRYVMOT 1200
OY 1141 TALDKALMBFMSKMEENKIKELMOOTYRGODIDYISINSDEGAGTSYRYVMOT 1200
OY 1201 GDAELENGRCSAGOKYLAIIIRLALAEFCLNGILALDEPTNLDGPAESLAALL 1260
OY 1201 GDAELENGRCSAGOKYLAIIIRLALAEFCLNGILALDEPTNLDGPAESLAALL 1260
OY 1261 RIMEARKQOENFOLIVITHDERFAHLIGORLAKRYRVSDENOHSTIESQETFD 1316
OY 1261 RIMEARKQOENFOLIVITHDERFAHLIGORLAKRYRVSDENOHSTIESQETFD 1316
OY 1261 RIMEARKQOENFOLIVITHDERFAHLIGORLAKRYRVSDENOHSTIESQETFD 1316

```

RESULT 2

AAW22775 standard; Protein; 1312 AA.

AAW22775;

21-DEC-1998 (first entry)

Human RAD50.

XX

```

KW Human; RAD50; DNA repair; tumour suppression; cancer; Septin-2;
KW central nervous system.
OS Homo sapiens.
PN WO9727284-A2.
XX 31-JUL-1997.
XX 24-JAN-1997; 97WO-US01299.
XX 17-JUL-1996; 96US-0687080.
XX 26-JAN-1996; 96US-0592126.
XX (GENE-) GENELABS TECHNOLOGIES INC.
XX Dolganov G;
XX WPI: 1997-393672/36.
XX N-PSDB; AAT75237.
XX Human tumour suppressor gene RAD50 - useful to detect
XX predisposition to, decrease risk of and treat cancer, also Septin-2
XX homologues
XX Claim 5; Page 82-86; 195pp; English.
XX The human RAD50 (hRAD50) is involved in DNA repair and has tumour
XX suppression activity, can be used to detect predisposition to, decrease
XX the risk of or treat cancers, e.g. acute myeloid leukaemia,
XX myelodysplastic syndrome, therapy related myelodysplastic syndrome,
XX therapy related acute myeloid leukaemia, refractory anaemia or refractory
XX anaemia with excess blasts. Also disclosed in this invention is Human
XX Septin-2 homologues of which may be used as targets for cancer therapies
XX and central nervous system directed treatment methods, and to measure the
XX proliferative potential of selected cell types.
XX Sequence 1312 AA;
XX
XX Query Match 2.4%; Score 31; DB 18; Length 1312;
XX Best local Similarity 100.0%; Pred. No. 5.7e-21;
XX Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1207 MRGRCSAGOKYLAIIIRLALAEFCLNGI 1237
OY 1197 MRGRCSAGOKYLAIIIRLALAEFCLNGI 1227
Db
RESULT 3
AAW71295
ID AAW71295 standard; Protein; 1312 AA.
XX
XX AAW71295;
XX
XX 25-NOV-1998 (first entry)
XX
XX Human homologue of yeast RAD50.
XX
XX Human homologue; yeast RAD50; Drosophila Septin-2; Acyl-CoA synthetase;
XX immunomodulatory activity; identification; activated T-cell.
XX Homo sapiens.
XX OS
XX WO9838306-A1.
XX
XX 03-SEP-1998.
XX
XX 27-FEB-1997; 97WO-US03159.
XX
XX 27-FEB-1997; 97WO-US03159.
XX
XX (GENE-) GENELABS TECHNOLOGIES INC.
XX

```

PI Dolgancov G;

XX MPI: 1998-481207/41.
DR N-PSDB: AAV59979.

XX Novel human immunomodulatory poly:peptide(s) - have homology to the
PT yeast RAD50 or Drosophila Septin-2 proteins
XX

PS Disclosure: Pages 136-140; 155pp; English.

XX The present sequence represents a human homologue of the yeast
CC S. cerevisiae gene RAD50. The present sequence has 35% overall
CC homology to the yeast RAD50 gene, and is expressed in activated
CC T-cells, testis, foetal liver and heart tissues. The specification
CC also describes sequences encoding human homologues of the
CC yeast RAD50, the Drosophila Septin-2 and Acyl-CoA synthetase. The
CC proteins have immunomodulatory activity. The nucleic acids and
CC proteins can be used to identify activated T-cells in a sample
CC population. They can also be used to isolate and identify sequences
CC encoding other proteins or other compounds having immunomodulatory
XX activity.

XX Sequence 1312 AA;

Query Match 2.4%; Score 31; DB 19; Length 1312;
Best Local Similarity 100.0%; Pred. No. 5.7e-21;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1207 MRGRCSAGOKVLASLIIRLALAEFTFCINCIGI 1237
|||||

DB 1197 MRGRCSAGOKVLASLIIRLALAEFTFCINCIGI 1227
|||||

Search completed: December 26, 2002, 08:45:17
Job time : 70 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 26, 2002, 05:07:47 ; Search time 7537 Seconds

(without alignments)
17345.071 Million cell updates/sec

Title: US-09-538-396-1

Perfect score: 4492

Sequence: 1 aattcgacagatgagatcc.....atctagcatcaaaaaaaaaa 4492

Scoring table: OLIGO_NUC
Gapop 60.0, Gapext 60.0

Searched: 2054640 seqs, 14551402878 residues

Word size: 30

Total number of hits satisfying chosen parameters: 3

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Database:

GenEmbl: *
1: gb_da: *
2: gb_hlg: *
3: gb_in: *
4: gb_om: *
5: gb_ov: *
6: gb_pat: *
7: gb_ph: *
8: gb_pl: *
9: gb_pr: *
10: gb_ro: *
11: gb_sts: *
12: gb_sy: *
13: gb_un: *
14: gb_vl: *
15: em_ba: *
16: em_fun: *
17: em_hum: *
18: em_in: *
19: em_mu: *
20: em_om: *
21: em_or: *
22: em_ov: *
23: em_pat: *
24: em_ph: *
25: em_pl: *
26: em_ro: *
27: em_sts: *
28: em_un: *
29: em_vl: *
30: em_hlg_hum: *
31: em_hlg_inv: *
32: em_hlg_other: *
33: em_hlg_mus: *
34: em_hlg_pln: *
35: em_hlg_rtd: *
36: em_hlg_mam: *
37: em_hlg_vrt: *
38: em_sy: *
39: em_hlgo_hum: *
40: em_hlgo_mus: *
41: em_hlgo_other: *

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	4492	100.0	4492	2	AX047235 Sequence
2	38	0.8	124610	2	AP004086 Oryza sat
3	38	0.8	136073	2	AP004773 Oryza sat

ALIGNMENTS

RESULT 1
AX047235
LOCUS AX047235
DEFINITION Sequence 1 from Patent WO0068404.
ACCESSION AX047235
VERSION AX047235.1 GI:11876516
KEYWORDS
SOURCE Zea mays.
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 4492)
Mahajan, P.B. and Shl, J.
Maize rad50 orthologue and uses of thereof
Patent: WO 0068404-A 1 16-NOV-2000;
PIONEER HI-BRED INTERNATIONAL, INC. (US)

REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
Source
CDS
1. .4492
/organism="Zea mays"
/db_xref="taxon:4577"
292. .4242
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC18962.1"
/db_xref="GI:11876517"
/translation="MSTVDKMLIKGIRSPDPDNKAVITFEKPLTIVGNGAKTTII
ECKLSTGELPNSRSGHVFDPVAGETKGOIKLFTPAAGKDVICRSOLT
OKASKMEFKAESVLOTINPHGERVCISRCADMRREIPALMGVSKATLVNITPHO
DESNPFLQDPSTLKKRDFISATRYTALAEVTKLAKDQMOGKIKTKLKLNDTVK
DQAHKLRENIAQDQKSDASKQEQKIKCTEREIILQMTSLDELRLQGOIDIK
ATERSTILTQOHEKIALSEENEDTDEELMEQTKFEERIALETKISLVMDDEA
SYSSVLSKONSELTHEIGKLAADAHLMKHERDSDIKNCKHNLGAPVPEPTND
VAMNLTNRITKARLSLENLDLKKKSNEDOLVIMKHYIKINARVSEVDGQIOSRTES
MSGILRRRDKERDAAYELSKPLNSIDEREKMOLEVERKTALGERYDSIIS
OKTEVYSLEOKTKVLLREKDIINRNADERVKGLKDLSSKDLNIVENHDKI
KTVLRGNPFPEKDMKEINQAPWPDKEVNEKSKQEAQELKTQSVYTAARQLT
KLRRDQAKRRFDSKLGILQISANAVDEPFLQDANKRDEOKLEFANGKEML
APEHLLARKNHVCCERAFPEDEEFKQKQMONSSTAESKALMERSVAEELFO
OLDKLRTIYDAYVYKVEETIPLAEKLNHLDSEOKAQAPDILGVLAHYMDADAV
BALOPTDITDRVHEIQQVKEVELEALDSSGCVASLEIQTLELFLDRTDTL
IVEVDLDQHRMLNEDMSSAVQRMNAEEVKASSILERQKSEEVVLAEEREO
LIVERKLLSEESLDPLSEKESLQAEYNALQKDEYHQIARKEPQOEDLALREKQ
MKIKGLYDSKKNKKEIKELQGRHYLCSQSQSQAKQORISAEILNSKEFLQGOGLK
NIDDKLTKRRTADVQLRDLIESLEERLISGSLSAIFADLRKRSOERERENSENR
WQGLTSVQOSNTSKHROELKSQYDIERRYNOFLQKTTMANKDLDRTYTALDKA
LRFHSMKMEIKIKELMOOTYIRGODIYISINDSGAATRSYSTRVYQOTDAE
LHRGRCSAGQVYASLITRLALAEFCNGILADEPTNLDPNMSLAAALRI
MEARKQENFOLIVITHDERFAHLIGQRLAARRYRVSDEQHSIISSEIFD"

BASE COUNT 1537 a 831 c 1061 g 1063 t
ORIGIN

Query Match 100.0%; Score 4492; DB 6; Length 4492;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 4492; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Pred. No. is the number of results predicted by chance to have a

```
|||||
Db 1 AATTGGCAGAGATGATCCATTAGACCCATAGCGTACAAAAACCTTAAGAACCTTAAC 60
QY 61 CGGTACAAAACCTTAACAAACCTTAAGCGGCTGAAGCTCCAAAAAAGCGATTTCTC 120
Db 61 CGGTACAAAACCTTAACAAACCTTAAGCGGCTGAAGCTCCAAAAAAGCGATTTCTC 120
QY 121 CTGCACTGCCCCCTCTTTCTCTTTCTCTTCCAAATCTTTTTCACATACAGAGCGTAATGAT 180
Db 121 CTGCACTGCCCCCTCTTTCTCTTTCTCTTCCAAATCTTTTTCACATACAGAGCGTAATGAT 180
QY 181 AGAAGTTGATAGGAGATATACATCCGCAATCTAGGTTTGGGGGCAATCGCTCTGGCCAGAC 240
Db 181 AGAAGTTGATAGGAGATATACATCCGCAATCTAGGTTTGGGGGCAATCGCTCTGGCCAGAC 240
QY 241 TGGATCGGATGCAAGTCGTAAGAGGAGGACATGGGGCTCTGGGGGCAAGATGAGCACC 300
Db 241 TGGATCGGATGCAAGTCGTAAGAGGAGGACATGGGGCTCTGGGGGCAAGATGAGCACC 300
QY 301 GTTGACAAGATGCTGATCAAGGGGATTCGAGCTTCGATCCGCAATTAAGAAAGTTCATC 360
Db 301 GTTGACAAGATGCTGATCAAGGGGATTCGAGCTTCGATCCGCAATTAAGAAAGTTCATC 360
QY 361 ACCTTCTTCAAGCCGCTCACCCTCATCTGTGGCCCCCAAGCGTGGTGGCAAGACAGATC 420
Db 361 ACCTTCTTCAAGCCGCTCACCCTCATCTGTGGCCCCCAAGCGTGGTGGCAAGACAGATC 420
QY 421 ATGAGAGCCGTAAGCTTTCTTGACCGGAGAGTGCCTCCCAACTCCGCTCTGGCCAC 480
Db 421 ATGAGAGCCGTAAGCTTTCTTGACCGGAGAGTGCCTCCCAACTCCGCTCTGGCCAC 480
QY 481 ACCTTCTTCAAGCCGCTCACCCTCATCTGTGGCCCCCAAGCGTGGTGGCAAGACAGATC 540
Db 481 ACCTTCTTCAAGCCGCTCACCCTCATCTGTGGCCCCCAAGCGTGGTGGCAAGACAGATC 540
QY 541 CGGTTTAAGATGCAAGCAGGAAAGATGTTGTCATCCGGTCTTCCAGCTTAACCAA 600
Db 541 CGGTTTAAGATGCAAGCAGGAAAGATGTTGTCATCCGGTCTTCCAGCTTAACCAA 600
QY 601 AAGGATCAAAAGATGAGATTAAAGCAATTGAAGCGTCTCCAGACTATTAATCCACAC 660
Db 601 AAGGATCAAAAGATGAGATTAAAGCAATTGAAGCGTCTCCAGACTATTAATCCACAC 660
QY 661 ACAGGGGAGAAAGTCTGCTCAGTCAAGATGCTGATGATGATGAGATTCCTCC 720
Db 661 ACAGGGGAGAAAGTCTGCTCAGTCAAGATGCTGATGATGATGAGATTCCTCC 720
QY 721 TTAATGGGTGTTTCAAGAGCGCTACTGGAAGATGTTATATTGTTTCCACCAAGATGATCC 780
Db 721 TTAATGGGTGTTTCAAGAGCGCTACTGGAAGATGTTATATTGTTTCCACCAAGATGATCC 780
QY 781 AATTGGCCATTGCAAGACCCGTCACACTTAAGAAAGTTGAGATCTCTCTGCC 840
Db 781 AATTGGCCATTGCAAGACCCGTCACACTTAAGAAAGTTGAGATCTCTCTGCC 840
QY 841 ACACGCTATACAAAGCTCTTGAAGTCAATAAGAACTTCAACAGATCAATGGAAG 900
Db 841 ACACGCTATACAAAGCTCTTGAAGTCAATAAGAACTTCAACAGATCAATGGAAG 900
QY 901 ATCAAGACTTTTAAAGTGAAGAACTTCAAGACTTCAAGAAAGACCAACATTAAG 960
Db 901 ATCAAGACTTTTAAAGTGAAGAACTTCAAGACTTCAAGAAAGACCAACATTAAG 960
QY 961 CTGCGTGAAGATATTGCTCAAGATCAAGAAAGTCAAGATGCTCAAAATCTCAGATGAG 1020
Db 961 CTGCGTGAAGATATTGCTCAAGATCAAGAAAGTCAAGATGCTCAAAATCTCAGATGAG 1020
QY 1021 CAACCTAAGGAAAGATCTGTGTACCGAGAGAAATCTCTGCAATGGAACAACTTGG 1080
Db 1021 CAACCTAAGGAAAGATCTGTGTACCGAGAGAAATCTCTGCAATGGAACAACTTGG 1080
QY 1081 GATGAAGTGAAGACTTCAAGGAGCAAAATTGACATCAAGGCAAGAGAGATGATCTTA 1140
Db 1081 GATGAAGTGAAGACTTCAAGGAGCAAAATTGACATCAAGGCAAGAGAGATGATCTTA 1140

|||||
Db 1081 GATGAAGTGAAGACTTCAAGGAGCAAAATTGACATCAAGGCAAGAGAGATGATCTTA 1140
QY 1141 CTTCAGCAGACATTAAGAAAGCTTGTGCACTTCTGAGGAAATGAAGATACCGATGAG 1200
Db 1141 CTTCAGCAGACATTAAGAAAGCTTGTGCACTTCTGAGGAAATGAAGATACCGATGAG 1200
QY 1201 GAACATATGAGATGCAACAAATTTGAAGAAAGATTCGCTTACTAGAACAAATATC 1260
Db 1201 GAACATATGAGATGCAACAAATTTGAAGAAAGATTCGCTTACTAGAACAAATATC 1260
QY 1261 AGTAACTTGAAGATATGATGATGAAGCATCTTAATAGCTCCGTTCTGTCCAAACGA 1320
Db 1261 AGTAACTTGAAGATATGATGATGAAGCATCTTAATAGCTCCGTTCTGTCCAAACGA 1320
QY 1321 AATTCTGAATTAACATATGAATTTGGAAGCTCCAGGACAGACCTGATGCTACCTGACT 1380
Db 1321 AATTCTGAATTAACATATGAATTTGGAAGCTCCAGGACAGACCTGATGCTACCTGACT 1380
QY 1381 ATGAGCATGAGAGAGATGACATTAATAATATATGCACTAAACATAATCTTGGGCCG 1440
Db 1381 ATGAGCATGAGAGAGATGACATTAATAATATATGCACTAAACATAATCTTGGGCCG 1440
QY 1441 GTTCTGTAACATCCCTTTACGAATGATGTTGCTATGACCTTACAAACAGATTTAAAGC 1500
Db 1441 GTTCTGTAACATCCCTTTACGAATGATGTTGCTATGACCTTACAAACAGATTTAAAGC 1500
QY 1501 AGACTATCAAGCTTGAAGATGATTTGCTGATGAAGAAATCCATGAAAGATCACTTA 1560
Db 1501 AGACTATCAAGCTTGAAGATGATTTGCTGATGAAGAAATCCATGAAAGATCACTTA 1560
QY 1561 GATGTTTGTGAAACATATCTTTAAATTAATTAATGCTGCTACTCCGAAGTTGATGTCAG 1620
Db 1561 GATGTTTGTGAAACATATCTTTAAATTAATTAATGCTGCTACTCCGAAGTTGATGTCAG 1620
QY 1621 ATCAATCTAAGATTTGATTCATGTCAGGCAATTTAAGAGGAGAAAGATTAAGAGAAA 1680
Db 1621 ATCAATCTAAGATTTGATTCATGTCAGGCAATTTAAGAGGAGAAAGATTAAGAGAAA 1680
QY 1681 GAAGCGGATGCTGAGAGATGAGAGCTTCAAAATTTAATATCCGATGATGAGAGG 1740
Db 1681 GAAGCGGATGCTGAGAGATGAGAGCTTCAAAATTTAATATCCGATGATGAGAGG 1740
QY 1741 GAGAGACATATGCAAAATTTGAAGTGAAGAGAACACTTGGCTTGGAGAAAGACTAT 1800
Db 1741 GAGAGACATATGCAAAATTTGAAGTGAAGAGAACACTTGGCTTGGAGAAAGACTAT 1800
QY 1801 GATTCATTTAATAGTCAGAAACGAAACAGAACTATATATGTTGGAACAGAAATTAAGTG 1860
Db 1801 GATTCATTTAATAGTCAGAAACGAAACAGAACTATATATGTTGGAACAGAAATTAAGTG 1860
QY 1861 CTTCGGGGGAGAAAGATATATATTAATTAATGAATGCTGATGAAGAGTAAACTGGGTTG 1920
Db 1861 CTTCGGGGGAGAAAGATATATATTAATTAATGAATGCTGATGAAGAGTAAACTGGGTTG 1920
QY 1921 AAGAAGATGATGAGAAAGAGCAAGAGCAAGCTCAATGAGATGTTAATGAGATTAAG 1980
Db 1921 AAGAAGATGATGAGAAAGAGCAAGAGCAAGCTCAATGAGATGTTAATGAGATTAAG 1980
QY 1981 GATTAATCAAAAAGTACTTATAGGGGAGAGATCTCTTTGGAAGAGATGAGAAAGAG 2040
Db 1981 GATTAATCAAAAAGTACTTATAGGGGAGAGATCTCTTTGGAAGAGATGAGAAAGAG 2040
QY 2041 ATCAATCAAGCTTTTGGCTGTGAGCAAGAAATGCAATGAGTTAAGATCAAAATCCAG 2100
Db 2041 ATCAATCAAGCTTTTGGCTGTGAGCAAGAAATGCAATGAGTTAAGATCAAAATCCAG 2100
QY 2101 GAAGCAGAGCAAGACTTAATTTAATCTCAGAGCAAAAGTATGATGCTATGAGAAATTTG 2160
Db 2101 GAAGCAGAGCAAGACTTAATTTAATCTCAGAGCAAAAGTATGATGCTATGAGAAATTTG 2160
QY 2161 ACAAAATCTTGAAGATATGATGCAAAAAGAAATTCCTGAGATCGGAAATCTCAATCT 2220
Db 2161 ACAAAATCTTGAAGATATGATGCAAAAAGAAATTCCTGAGATCGGAAATCTCAATCT 2220
```


QY	3301	1GCAATGGCAAAACAGCAAAAGAAATATCAGCTGAGTTAAACAAGACAAAGAACTACTGCG	3360
Db	3301	TGCAATGGCAAAACAGCAAAAGAAATATCAGCTGAGTTAAACAAGACAAAGAACTACTGCG	3360
QY	3361	GGCCAGGGCCAGTTGGAAAAAGAACTTGAATGACATCTCAAGTACAGGAAAAACAAGGCT	3420
Db	3361	GGCCAGGGCCAGTTGGAAAAAGAACTTGAATGACATCTCAAGTACAGGAAAAACAAGGCT	3420
QY	3421	GATGTGGAACAACCTTACTCTCGATATAGAAATCACCTTGAAAGAAAGGCTGCTTCAATAGGT	3480
Db	3421	GATGTGGAACAACCTTACTCTCGATATAGAAATCACCTTGAAAGAAAGGCTGCTTCAATAGGT	3480
QY	3481	AGCTTGCTGCTATAGAAAGCTGATCTGAAAGCCCAATCTCAAGAAAAAGAGAGGCTTAAT	3540
Db	3481	AGCTTGCTGCTATAGAAAGCTGATCTGAAAGCCCAATCTCAAGAAAAAGAGAGGCTTAAT	3540
QY	3541	TCGAATTTAAACAGGTGGCAAGAACACTTCTGTTATCAAAGTAAATTTTCAAAGCAC	3600
Db	3541	TCGAATTTAAACAGGTGGCAAGAACACTTCTGTTATCAAAGTAAATTTTCAAAGCAC	3600
QY	3601	AAACAAGAGCTTAAACTGTCAACAGTAAAGGATTCGAGAAGGATATCTACTAATCAATTT	3660
Db	3601	AAACAAGAGCTTAAACTGTCAACAGTAAAGGATTCGAGAAGGATATCTACTAATCAATTT	3660
QY	3661	CTCCAGCTTAAGACAACACTGAAATGGCAACAGGACTTGGACAGATATTTACTGCTTTA	3720
Db	3661	CTCCAGCTTAAGACAACACTGAAATGGCAACAGGACTTGGACAGATATTTACTGCTTTA	3720
QY	3721	GACAAGGCTCTTATGGCGGTTCCACAGATGAAGTGGAGAGATAAATTAATTAATCAG	3780
Db	3721	GACAAGGCTCTTATGGCGGTTCCACAGATGAAGTGGAGAGATAAATTAATTAATCAG	3780
QY	3781	GAACTGGGCAACAGACATACAGAGGCGCAGGATTTGATTACATTAAGCATTAATTTCTGAT	3840
Db	3781	GAACTGGGCAACAGACATACAGAGGCGCAGGATTTGATTACATTAAGCATTAATTTCTGAT	3840
QY	3841	TCTGAGGGTGTGGCACTCGATCATACAGCTACCGGTTGTTATGCAAACTGGTGAATGCT	3900
Db	3841	TCTGAGGGTGTGGCACTCGATCATACAGCTACCGGTTGTTATGCAAACTGGTGAATGCT	3900
QY	3901	GACCTGGAAATGGGAAGGGGCTGCAATGCTGGTGCACAAAGTTCTTCTTTATAAATC	3960
Db	3901	GACCTGGAAATGGGAAGGGGCTGCAATGCTGGTGCACAAAGTTCTTCTTTATAAATC	3960
QY	3961	AGACTGAGCACTTCGGGAACTTTCTGCGTGAATCGGGATATATGCTTTGGATGAGCCA	4020
Db	3961	AGACTGAGCACTTCGGGAACTTTCTGCGTGAATCGGGATATATGCTTTGGATGAGCCA	4020
QY	4021	ACTACGAATCTAGATGGGCCAAATGACAGAGAGTCTTGCTGCGCTGTTGGAATTAATG	4080
Db	4021	ACTACGAATCTAGATGGGCCAAATGACAGAGAGTCTTGCTGCGCTGTTGGAATTAATG	4080
QY	4081	GAAAGCAGGAAAGGGCAGAGAACTTCCAGTTGATTTGTAATCACTCATGATGAGAGATTT	4140
Db	4081	GAAAGCAGGAAAGGGCAGAGAACTTCCAGTTGATTTGTAATCACTCATGATGAGAGATTT	4140
QY	4141	GGCCATCTTATGGTCAAAAGGAGAGCTGCTGGAAGTACTATACGATCTCCAAAGATGAG	4200
Db	4141	GGCCATCTTATGGTCAAAAGGAGAGCTGCTGGAAGTACTATACGATCTCCAAAGATGAG	4200
QY	4201	AAACAAGCAGAGATTAATGGAATCCCAAGAGATATTTGACTAAAGGGTGTTCTAGAGGCTG	4260
Db	4201	AAACAAGCAGAGATTAATGGAATCCCAAGAGATATTTGACTAAAGGGTGTTCTAGAGGCTG	4260
QY	4261	TGACAGCAGACCTGTTGCTAGTCGAATCCAGTAAATTTATGCGCAAGTACTGCTGCCAAG	4320
Db	4261	TGACAGCAGACCTGTTGCTAGTCGAATCCAGTAAATTTATGCGCAAGTACTGCTGCCAAG	4320
QY	4321	CAATGTTAAACAAGCTTTAGAGAGCTCTGTACGCCGTTACAGCAGTTCGCGGAAACCTAT	4380
Db	4321	CAATGTTAAACAAGCTTTAGAGAGCTCTGTACGCCGTTACAGCAGTTCGCGGAAACCTAT	4380
QY	4381	CCCTGCTGTTGTAATTAATTTAATCTGGCCAGAGATGGAATGTGTGCACTGGGTGAT	4440

DB 4381 CCTCGTGTGTAATTAATCTGCGCAGAGATGATGTCTCAGCTGAT 4440
 4441 GGATGTTTCACGACATCATGATGTTTCAATCTAGCATCAAAAAA 4492
 4442
 DB 4441 GGATGTTTCACGACATCATGATGTTTCAATCTAGCATCAAAAAA 4492

RESULT 2
 AP004086/c 124610 bp DNA linear HTG 21-MAR-2002

LOCUS Oriza sativa (japonica cultivar-group) chromosome 2 clone
 CUI288.D09, *** SEQUENCING IN PROGRESS ***, in ordered pieces.

ACCESSION AP004086
 VERSION AP004086.1 GI:15281365
 KEYWORDS HTG, PHASE2.

ORGANISM Oriza sativa (japonica cultivar-group) (cultivar:Nipponbare) DNA,
 clone:CUI288.D09.

REFERENCE 1
 AUTHORS Sasaki,T., Matsumoto,T. and Yamamoto,K.
 TITLE Oriza sativa nippobare(GA3) genomic DNA, chromosome 2, BAC
 clone:CUI288.D09
 JOURNAL Published only in Database (2001)
 REFERENCE 2 (bases 1 to 124610)
 AUTHORS Sasaki,T., Matsumoto,T. and Yamamoto,K.

COMMENT
 The nucleotide sequence of this BAC clone was generated by
 combining Monsanto and RGP-Japan sequencing data.
 NOTE: It currently consists of 1 contigs. Gaps between the contigs
 are represented as runs of N. The order of the pieces is believed
 to be correct as given, however the sizes of the gaps between them
 are based on estimates that have provided by the submitter. This
 sequence will be replaced by the finished sequence as soon as it is
 available and the accession number will be preserved.
 * NOTE: This is a 'working draft' sequence.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.

FEATURES
 source

1. 124610
 /organism="Oriza sativa (japonica cultivar-group)"
 /cultivar="Nipponbare"
 /db_xref="taxon:39947"
 /chromosome="2"
 /clone="CUI288.D09"

BASE COUNT 35712 a 27346 c 25887 g 35511 t 154 others
 ORIGIN

Query Match 0.8% Score 38; DB 2; Length 124610;
 Best Local Similarity 100.0%; Pred. No. 4.4e-08;
 Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3892 GGTGATGCTGAGCTGGAATGCGAGGCGCTGCACTGC 3929
 61062 GGTGATGCTGAGCTGGAATGCGAGGCGCTGCACTGC 61025

RESULT 3
 AP004773 136073 bp DNA linear HTG 21-FEB-2002
 LOCUS Oriza sativa (japonica cultivar-group) chromosome 2 clone P0416F07,
 DEFINITION *** SEQUENCING IN PROGRESS ***, in ordered pieces.
 ACCESSION AP004773
 VERSION AP004773.1 GI:18844993

KEYWORDS HTG, PHASE2.
 SOURCE Oriza sativa (japonica cultivar-group) (cultivar:Nipponbare) DNA,
 clone:P0416F07

ORGANISM Oriza sativa (japonica cultivar-group)
 clone:P0416F07
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE 1
 AUTHORS Sasaki,T., Matsumoto,T. and Yamamoto,K.
 TITLE Oriza sativa nippobare(GA3) genomic DNA, chromosome 2, PAC
 clone:P0416F07
 JOURNAL Published only in Database (2002)
 REFERENCE 2 (bases 1 to 136073)
 AUTHORS Sasaki,T., Matsumoto,T. and Yamamoto,K.

COMMENT
 The nucleotide sequence of this BAC clone was generated by
 combining Monsanto and RGP-Japan sequencing data.
 NOTE: It currently consists of 1 contigs. Gaps between the contigs
 are represented as runs of N. The order of the pieces is believed
 to be correct as given, however the sizes of the gaps between them
 are based on estimates that have provided by the submitter. This
 sequence will be replaced by the finished sequence as soon as it is
 available and the accession number will be preserved.
 * NOTE: This is a 'working draft' sequence.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.

FEATURES
 source

1. 136073
 /organism="Oriza sativa (japonica cultivar-group)"
 /cultivar="Nipponbare"
 /db_xref="taxon:39947"
 /chromosome="2"
 /clone="P0416F07"

BASE COUNT 38865 a 28336 c 29834 g 38561 t 477 others
 ORIGIN

Query Match 0.8% Score 38; DB 2; Length 136073;
 Best Local Similarity 100.0%; Pred. No. 4.4e-08;
 Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3892 GGTGATGCTGAGCTGGAATGCGAGGCGCTGCACTGC 3929
 108120 GGTGATGCTGAGCTGGAATGCGAGGCGCTGCACTGC 108157

Search completed: December 26, 2002, 07:46:52
 Job time : 8293 secs